

Article

Energy-Related CO₂ Emissions Forecasting Using an Improved LSSVM Model Optimized by Whale Optimization Algorithm

Haoran Zhao, Sen Guo * and Huiru Zhao

School of Economics and Management, North China Electric Power University, Beijing 102206, China; 1162106030@ncepu.edu.cn (H.Z.); zhaohuiru@ncepu.edu.cn (H.Z.)

* Correspondence: guosen@ncepu.edu.cn or guosen324@163.com; Tel.: +86-010-6177-3134

Academic Editor: Pierre Trambouze

Received: 31 May 2017; Accepted: 26 June 2017; Published: 29 June 2017

Abstract: Accurate and reliable forecasting on energy-related carbon dioxide (CO₂) emissions is of great significance for climate policy decision making and energy planning. Due to the complicated nonlinear relationships of CO₂ emissions with its driving forces, the accurate forecasting for CO₂ emissions is a tedious work, which is an important issue worth studying. In this study, a novel CO₂ emissions prediction method is proposed which employs the latest nature-enlightened optimization method, named the Whale optimization algorithm (WOA), to search the optimized values of two parameters of LSSVM (least squares support vector machine), namely the WOA-LSSVM model. Meanwhile, the driving forces of CO₂ emissions including GDP (gross domestic product), energy consumption and population are chosen to be the import variables of the proposed WOA-LSSVM method. Taking China's CO₂ emissions as an instance, the effectiveness of WOA-LSSVM-based CO₂ emissions forecasting is verified. The comparative analysis results indicate that the WOA-LSSVM model is significantly superior to other selected models, namely FOA (fruit fly optimization algorithm)-LSSVM, LSSVM, and OLS (ordinary least square) models in terms of CO₂ emissions forecasting. The proposed WOA-LSSVM model has the potential to effectively improve the accuracy of CO₂ emissions forecasting. Meanwhile, as a new nature-enlightened heuristic optimization algorithm, the WOA has the prospect for wide application.

Keywords: CO₂ emissions prediction; Whale optimization algorithm; LSSVM; parameters optimization; driving forces

1. Introduction

With the high speed expansion of globalization and industrialization, energy consumption of the whole world has experienced a rapid increase in the last 2 decades. The consumption of fossil fuel, contributing to economic development to a large extent, constitutes 80% of the global energy consumption [1]. It has become widely that greenhouse gases emissions, particularly CO₂ emissions, have destructive impacts on the environment, especially in terms of global warming, primarily from such sources as fossil fuel burning for electricity production and heat supply [2]. In the past 20 years, nearly 75% of the human-caused carbon emissions originated from fossil fuel burning. The scientific opinion holds that CO₂ emissions are the main driving forces of climate change [3]. Therefore, it is essential to forecast CO₂ emissions from the angle of energy consumption for climate policy making and energy planning. First, CO₂ emissions forecasting can not only help predict how future global temperatures will rise, but also help to evaluate prospective expenses of emission reduction, as well as potential benefits from guarding against global temperatures rising [4]. Second, the Copenhagen Accord established that global temperatures rising ought to be confined to below 2 °C, which demands

40–70% emissions decrease by 2050, as compared to emissions in 2010. With the approaching of 2050, various carbon capture technologies have been developed for CO₂ emissions reduction to achieve a 2 °C limitation on global warming [5–9]. In addition to carbon capture technologies, CO₂ emissions prediction can also contribute toward examining whether the sustained reinforcing of global climate policy has the capability to bring a notable CO₂ emissions decrease to achieve the 2 °C goal. If not, the governments need to carry out more aggressive climate policies.

Several kinds of estimation methods have been developed in CO₂ forecasting modeling in recent years. The structure model, a popular approach, has been employed in some literatures to forecast CO₂ emissions, such as the national energy modeling system (NEMS) and computable general equilibrium (CGE) models [10–12]. Nevertheless, these approaches are unable to reflect the dynamic instinct of the real economy, due to the stationary of parameter values in structural models. Therefore, there are also other methods used to forecast CO₂ emissions. Meng [13] employed a logistic function to imitate emissions generated from non-renewable energy burning. Liang [14] constructed an input–output model of energy consumption and CO₂ emissions of multi-areas in China and carried out scenario analysis for 2010 and 2020. Chen put forward a blended fuzzy regression and back propagation network method, which was used to forecast global CO₂ concentration [15]. Pao [16] and Lin [17] employed grey model to forecast CO₂ emissions of Brazil and Taiwan. Through estimating China's future energy demand, He [18] predicted CO₂ emissions during the period of 2010 to 2020 on the basis of scenario analysis method. From the previous studies, it can be seen that there are few studies employing the intelligent forecasting techniques (algorithms), which usually have the advantages of strong learning ability and high forecasting accuracy. Therefore, this paper attempts to employ the LSSVM (least squares support vector machine) technique, a famous and widely applied intelligent algorithm, to forecast CO₂ emissions.

The LSSVM model has been widely utilized to deal with forecasting issues in different fields, such as gas [19], short-term electric load [20], wind speed [21], and hydropower consumption [22]. However, it is rare to find that the LSSVM technique has been used for CO₂ emissions forecasting. Therefore, the application feasibility of LSSVM model to forecast CO₂ emissions is examined in this paper. Generally speaking, the LSSVM method relies on the values of two parameters, namely the bandwidth of RBF kernel " σ " and the regularization parameter " c ". At present, some optimization methods were utilized to find the optimized values for the two parameters, such as genetic algorithm [20], artificial bee colony algorithm [23], chaotic differential evolution approach [24], and FOA (fruit fly optimization algorithm) [25]. However, most of these optimization methods have some disadvantages of being hard to understand, or achieving local optimum but reaching global optimum slowly. The WOA (Whale optimization algorithm), put forward by Seyedali Mirjalili in 2016, is a novel heuristic optimization algorithm enlightened by nature [26]. This new optimization method can be easily understood and achieve the global optimal solution in a quite short time. Considering these superiorities, this study tries to employ the WOA method to optimize two significant parameters of the LSSVM method, with the aim of improving the forecasting performance of CO₂ emissions. Simultaneously, at the aim of further improving the prediction accuracy, the main driving factors of CO₂ emissions including GDP (gross domestic product), energy consumption and population, are also taken to be the import variables for the established WOA-LSSVM approach to forecast CO₂ emissions.

The primary contributions of this paper are the following:

- A novel LSSVM-based CO₂ emissions forecasting model is proposed, of which the significant parameters in LSSVM are optimized by the meta-heuristic optimization algorithm WOA. It is verified that this hybrid intelligence forecasting method has the superiority in the forecasting precision of CO₂ emissions through comparing with LSSVM model optimized by FOA, LSSVM without parameters optimization and OLS (ordinary least square). This paper extends the application domains of intelligence LSSVM forecasting technique.
- GDP, energy consumption and population, which are considered as the main driving forces of CO₂ emissions, are imported into the proposed WOA-LSSVM approach. Therefore, the proposed CO₂

emissions prediction model in this paper not only employs the intelligence forecasting technique, but also takes the social economic driving factors of CO₂ emissions into consideration, which encapsulates the complicated nonlinear relationships of CO₂ emissions with its main driving forces to some extent.

The remainder of the study is processed below. Section 2 represents the theory of WOA method and LSSVM model, and a novel CO₂ emissions prediction method (namely WOA-LSSVM) that combines the WOA and LSSVM technique is illustrated. Section 3 proceeds with the empirical simulation and analysis for WOA-LSSVM-based CO₂ emissions. In Section 4, the CO₂ emissions forecasting performance of different models are compared. Finally, the conclusions are presented in Section 5.

2. The Methodology of WOA-LSSVM Forecasting Model

2.1. Basic Methodology of LSSVM Model

LSSVM is an improvement of the support vector machine (SVM), which replaces the inequality constraints of traditional SVM with equality constraints, studying with the principle of structural risk minimization through employing linear least squares guide lines to the loss function. So, the speed of problem solving and the accuracy of convergence can be enhanced [27]. The fundamental theory of LSSVM method is introduced as below.

Set a series of samples $\{x_i, y_i\}_{i=1}^m$, taking $x_i \in R_n$ as the entering vector and $y_i \in R$ as the homologous output value of sample i . Through employing a nonlinear function ϕ , the sample data are mapped to a higher dimension from the customary feature space, therefore, to approximately describe it in a linear form:

$$f(x) = w^T \phi(x) + b \quad (1)$$

where w implies the weight vector, and b indicates the error.

In the original space, the LSSVM with the equality constraints is expressed as:

$$\begin{cases} \min J(w, \xi) = \frac{1}{2} w^T w + \frac{1}{2} C \sum_{i=1}^m \xi_i^2 \\ s.t. y = w \phi(x) + b + \xi_i \end{cases} \quad (2)$$

where C indicates the regularization parameter, and ξ_i represents the slack variable.

Then, the Lagrangian function L can be established as:

$$L(w, b, \xi, a) = \frac{1}{2} w^T w + \frac{1}{2} C \sum_{i=1}^m \xi_i^2 - \sum_{i=1}^m a_i \{ w^T \phi(x_i) + b + \xi_i - y_i \} \quad (3)$$

where a_i implies Lagrange multiplier.

The conditions of Karush–Kuhn–Tucker (KKT) for optimal performance are provided by:

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^m a_i \phi(x_i) \\ \frac{\partial L}{\partial b} = 0 \rightarrow \sum_{j=1}^m a_j = 0 \\ \frac{\partial L}{\partial \xi_i} = 0 \rightarrow a_i = C \xi_i \\ \frac{\partial L}{\partial a_i} = 0 \rightarrow w^T \phi(x_i) + b + \xi_i - y_i \end{cases} \quad (4)$$

Removing the variables w and ξ_i , the majorization issue is converted into the linear form as follows:

$$\begin{bmatrix} 0 & Q^T \\ Q & K + C^{-1}I \end{bmatrix} \begin{bmatrix} b \\ A \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \quad (5)$$

where $Q = [1, \dots, 1]^T$, $A = [a_1, a_2, \dots, a_m]^T$; $Y = [y_1, y_2, \dots, y_m]^T$. In accordance with the Mercer's condition, the Kernel function is described as:

$$K(x_i, x_j) = \phi(x_i)^T \phi(x_j) \quad (6)$$

The LSSVM method for regression can be set by:

$$f(x) = \sum_{i=1}^m a_i K(x, x_i) + b \quad (7)$$

Owing to a fewer parameters to be set and a superior performance, the radial basis function (RBF) is chosen to be the Mercer kernel function $K(x, x_i)$ in this paper, which is shown in Equation (8).

$$K(x, x_i) = \exp\left\{-\|x - x_i\|^2 / 2\sigma^2\right\} \quad (8)$$

Consequently, two parameters are needed to be found if the LSSVM model is utilized to forecast CO₂ emissions, which are the bandwidth of RBF kernel " σ " and the regularization parameter " c ". The latest intelligent algorithm WOA will be applied to search the optimal values of " σ " and " c ", and the optimization details are introduced in Sections 2.2 and 2.3.

2.2. Basic Theory of WOA

Enlightened by the humpback whales' special hunting strategy, named the bubble-net feeding method [28], the researcher Seyedali Mirjalili proposed a novel nature-enlightened heuristic optimization method, namely WOA, simulating the public behavior of humpback whales [26]. The foraging behavior investigated by Goldbogen et al. employing tag sensors has found that there are two tactics bound up with the bubble, which are called "upward-spirals" and "double-loops" [29]. In "upward-spirals" strategy, humpback whales dive about 12 m and begin to generate bubbles in a heliciform around the prey and swim up to the surface, which is shown in Figure 1, taken from the original in study [26]. Therefore, it is worth mathematically imitating the heliciform bubble-net predation strategy to carry out optimization. WOA is proposed in accordance with this behavior.



Figure 1. Bubble-net predation strategy of humpback whales [26].

For a detailed description of WOA, please refer to the work documented in study [26]. In accordance with the upward spiral bubble-net foraging maneuver, the WOA method is made up of several steps, which are illustrated as follows:

- Step 1: Parameters setting.

The vital parameters of WOA contain: the amount of search agents $SearchAgents_no$; the amount of variables dim ; the maximum amount of generations $Max_iteration$; the lower bound $lb = [lb_1, lb_2, lb_3, \dots, lb_{n-1}, lb_n]$, and upper bound $ub = [ub_1, ub_2, ub_3, \dots, ub_{n-1}, ub_n]$ of variables.

- Step 2: Population initialization.

As a method on the basis of population, the humpback whales in WOA method can be expressed in a matrix as below:

$$W = \begin{bmatrix} w_{1,1} & w_{1,2} & \dots & \dots & w_{1,d} \\ w_{2,1} & w_{2,2} & \dots & \dots & w_{2,d} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ w_{n,1} & w_{n,2} & \dots & \dots & w_{n,d} \end{bmatrix} \quad (9)$$

where W represents the position matrix of humpback whales; w_{ij} means the value of j -th parameter of the i -th humpback whale, $i = 1, 2, \dots, n, j = 1, 2, \dots, d$, w_{ij} can be calculated by employing stochastic distribution which is indicated in Equation (10).

$$W(i, j) = (ub(i) - lb(i)) \times rand(i, j) + lb(i) \quad (10)$$

where $W(i, j)$ is the value of the i -th row, j -th column of the matrix; $lb(i)$ and $ub(i)$ imply the lower bound and upper bound of i -th humpback whale; $rand(i, j)$ represents the stochastic number generated from uniform distribution in $[0, 1]$ interval.

- Step 3: Fitness function determination.

At the aim of evaluating each humpback whale, the fitness function $f[*]$ ought to be determined during the process of majorization, and the matrix OW is utilized to store the fitness values of humpback whales as follows:

$$OW = \begin{bmatrix} OW_1 \\ OW_2 \\ \vdots \\ OW_n \end{bmatrix} = \left\{ \begin{array}{l} f[(w_{11} \ w_{12} \ \dots \ w_{1,d})] \\ f[(w_{21} \ w_{22} \ \dots \ w_{2,d})] \\ \vdots \\ f[(w_{n1} \ w_{n2} \ \dots \ w_{n,d})] \end{array} \right\} \quad (11)$$

- Step 4: Iteration process.

In WOA method, the search agents renovate their places according to a stochastically selected search agent or the best scheme attained till now. At the aim of performing a global search and avoiding local optimum, the WOA method begins with a series of random solutions, which means the location of a search agent would be updated with regard to a stochastically selected search agent. The mathematical model can be estimated as below:

$$\vec{D} = |\vec{C} \cdot \vec{X}_{rand} - \vec{X}| \quad (12)$$

$$\vec{X}(t+1) = \vec{X}_{rand} - \vec{A} \cdot \vec{D} \quad (13)$$

where t represents the current iteration, \vec{A} and \vec{C} mean coefficient vectors, \vec{X}_{rand} represents a stochastic location vector (a stochastic humpback whale) selected from the present population, \vec{X} implies the location vector, $||$ stands for the absolute value, and \cdot indicates an element-be-element multiplication. The vectors \vec{A} and \vec{C} can be computed as below:

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a} \quad (14)$$

$$\vec{C} = 2 \cdot \vec{r} \quad (15)$$

where \vec{a} is linearly reduced from 2 to 0 through the process of iterations, and \vec{r} represents a stochastic vector in $[0, 1]$.

Since the value of \vec{A} depends on the diversification of \vec{a} , a stochastic search agent is selected when $|\vec{A}| > 1$, while the best scheme is chosen when $|\vec{A}| < 1$ to renovate the location of the search agents. Under the condition of $|\vec{A}| < 1$, the humpback whales swim around the prey in two forms of paths, which are the shrinking circle and heliciform path.

For the shrinking encircling path, the search agents can renew their positions on the basis of the following equations:

$$\vec{D} = \left| \vec{C} \cdot \vec{X}^*(t) - \vec{X}(t) \right| \quad (16)$$

$$\vec{X}(t+1) = \vec{X}^*(t) - \vec{A} \cdot \vec{D} \quad (17)$$

where X^* indicates the location vector of the best solution attained till now. The vectors \vec{A} and \vec{C} can be calculated in accordance with Equations (14) and (15). Through reducing the value of \vec{a} in Equation (14), the shrinking encircling mechanism can be achieved.

For the spiral shaped path, the helix-shaped movement between the humpback whales and prey can be simulated as follows:

$$\vec{X}(t+1) = \vec{D}' \cdot e^{bl} \cdot \cos(2\pi l) + \vec{X}^*(t) \quad (18)$$

where $\vec{D}' = \left| \vec{X}^*(t) - \vec{X}(t) \right|$ implies the distance of the i -th humpback whale to the prey (best scheme attained till now), b means a constant used to determine the shape of the logarithmic spiral, l indicates a stochastic number in the interval $[-1, 1]$, and \cdot illustrates an element multiplication.

Since the humpback whales swim around the best solution within a shrinking circle path and along a heliciform path at the same time, for the purpose of modeling this simultaneous behavior, it is assumed that there is a 50% probability of selecting either the spiral model or the shrinking circle mechanism to renovate the positions of humpback whales in the majorization process. The mathematical equation can be established as below:

$$\vec{X}(t+1) = \begin{cases} \vec{X}^*(t) - \vec{A} \cdot \vec{D} & \text{if } p < 0.5 \\ \vec{D}' \cdot e^{bl} \cdot \cos(2\pi l) + \vec{X}^*(t) & \text{if } p \geq 0.5 \end{cases} \quad (19)$$

where p represents a stochastic number in interval $[0, 1]$.

- Step 5: Optimal selection.

During every iteration process, the search agents renovate their places with regard to a stochastically selected search agent or the best scheme attained till now. A stochastic search agent is selected under the condition of $|\vec{A}| > 1$ and the search agents renew their locations in accordance with the Equation (13), while the best solution is chosen if $|\vec{A}| < 1$ and the search agents renovate their places according to Equations (17) and (18) relying on the value of p . Finally, the WOA method comes to the termination if the iteration criterion is satisfied.

The primary steps of WOA method are demonstrated in Figure 2.

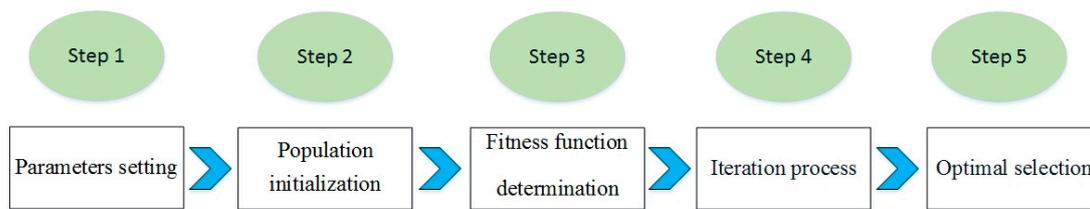


Figure 2. The steps of WOA method.

2.3. Basic Principle of WOA-LSSVM Model for CO₂ Emissions Forecasting

It is necessary to determine the optimal values of two parameters for LSSVM method before the operation of CO₂ emissions prediction. For the purpose of improving the forecasting performance, the WOA method is utilized to optimize the value of these two parameters.

The process of WOA-LSSVM model for CO₂ emissions forecasting is elaborated on below:

- Step 1: Parameters setting.

There are five primary parameters of WOA needed to be set, including the number of whales *SearchAgents_no*, the number of variables *dim*, the maximum number of iteration *Max_iteration*, the lower bound $lb = [lb_1, lb_2, lb_3, \dots, lb_{n-1}, lb_n]$, and upper bound $ub = [ub_1, ub_2, ub_3, \dots, ub_{n-1}, ub_n]$. In this paper, set *SearchAgents_no* = 50, *dim* = 2, *Max_iteration* = 100, $lb = [0.01, 0.001]$, and $ub = [100, 10]$.

- Step 2: Population initialization.

Since the values of *SearchAgents_no*, *dim*, *Max_iteration*, the lower bound and the upper bound have been initialized, the first stochastic population (location) of whales can be calculated using Equation (10). And the value of iteration is 1 at the original stage.

- Step 3: Fitness function determination.

In this study, the positions of whales are used to represent two parameters of LSSVM model, namely $w_{i,1} = \sigma$, and $w_{i,2} = c$. Therefore, $w_{i,1}$ and $w_{i,2}$ are fed into the LSSVM model for CO₂ emissions prediction. In accordance with the CO₂ emissions prediction result, the homologous value of fitness function can be computed based on the root mean square error (RMSE) shown in Equation (20) utilized to establish the fitness function in this paper.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{k=1}^n (x(k) - \hat{x}(k))^2} \quad (20)$$

where $x(k)$ indicates the actual value of CO₂ emissions at time k ; $\hat{x}(k)$ implies the forecasting value of CO₂ emissions at time k .

- Step 4: Optimization begins.

At the original iteration (iteration = 1), the fitness values of all the whales are computed based on Equation (20). Then, rank the first population of whales according to their fitness values, and choose the whale with the optimal fitness value. After the best whale is identified, other whales will attempt to renew their places with respect to the best whale, according to Equation (19).

- Step 5: Optimization ends.

When the first iteration comes to the termination, the best whale and the best fitness value of the whale can be obtained. Then, we can begin to generate the offspring generation, and renovate the location and fitness of the whale at each iteration on the basis of Equations (9)–(19). The whales update their positions with regard to the best whales selected according to the fitness value. The best whale and their corresponding fitness value can be obtained after each iteration. Therefore, after 100

iterations, the optimization process comes to the end, and the best position of whale can be obtained. Simultaneously, the regularization parameter “ c ” and the bandwidth of RBF kernel “ σ ” of LSSVM method are attained.

The optimization process of WOA-LSSVM method for CO₂ emissions prediction is illustrated in Figure 3.

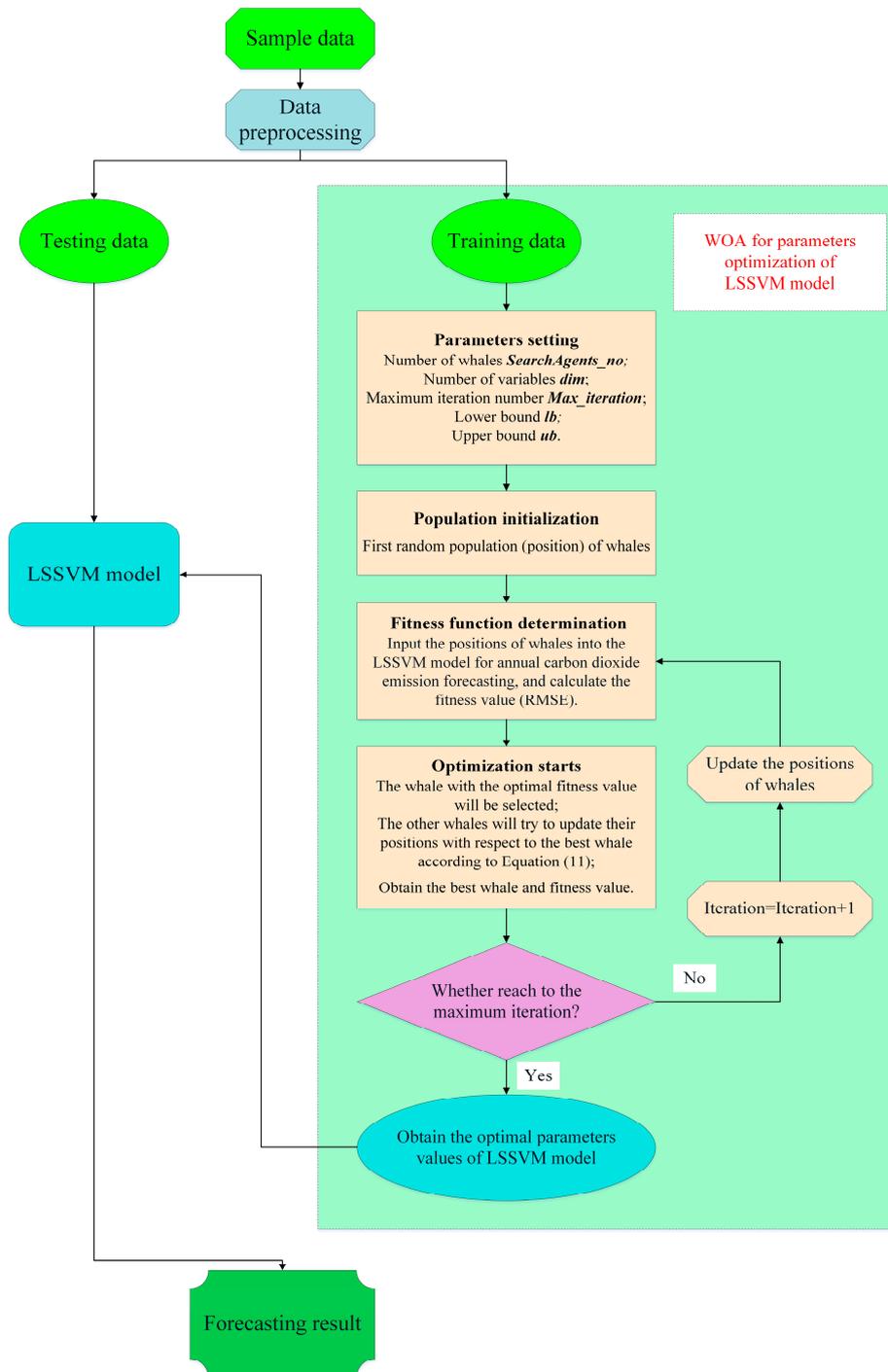


Figure 3. The procedure of the whale optimization algorithm of the least squares support vector machine (WOA-LSSVM) method for carbon dioxide (CO₂) emissions prediction.

3. Empirical Simulation and Analysis

3.1. Data Sources and Preprocessing of Data Sample

In this study, the WOA-LSSVM method is utilized to predict the CO₂ emissions in China. Before forecasting CO₂ emissions, it is critical to determine the key drivers of China's increasing CO₂ emissions, which should be considered as the input variables of the proposed WOA-LSSVM-based CO₂ emissions forecasting model. The CO₂ emissions and its driving forces show complicated nonlinear relationships. In summarizing previous studies [30–32], GDP, population, total exports, energy consumption, and economic structure are found to be the main drivers of CO₂ emissions. Based on these findings, we further explore the correlation degrees between these driving forces and China's CO₂ emissions between 1991 and 2014 by using the grey correlation analysis method, which measures the correlation degree between different factors according to the similarity or difference degree of trends, and the values are listed in Table 1. From Table 1, we can infer that GDP, energy consumption and population maintains the top three highest correlation degrees with China's CO₂ emissions, which are 0.6577, 0.6727, and 0.6798, respectively. Therefore, GDP, energy consumption and population are selected as the input variables for WOA-LSSVM-based CO₂ emissions forecasting model in this paper.

Table 1. The correlation degrees between China's CO₂ emissions and driving forces.

| Driving Forces | GDP | Energy Consumption | Population | Total Export | Economic Structure * |
|--------------------|--------|--------------------|------------|--------------|----------------------|
| Correlation degree | 0.6577 | 0.6727 | 0.6798 | 0.6377 | 0.6015 |

* The economic structure is represented by the proportion of the secondary industry added value in GDP.

The data of China's CO₂ emissions related to energy consumption from 1991 to 2014 are picked out from the British Petroleum (BP) Statistical Review of World Energy, which presents objective and high-quality data for world energy markets, including data about petroleum, coal, solar energy, natural gas, hydropower, wind power and nuclear power. According to the CO₂ emissions data of 68 countries for the period from 1965 to 2014 provided in BP Statistical Review of World Energy, the CO₂ emissions in China has a 5.58% average annual growth rate. As the largest CO₂ emitter, China emitted 9.76 Gt CO₂ in 2014, which increased 0.90% in comparison to that in 2013, which is a little more than the 0.5% global mean annual growth rate. The data of GDP, energy consumption, and population from 1991 to 2014 were picked out from the China Statistical Yearbook. Data on GDP were converted to the constant price by taking 2000 as the basic period.

The forecasting procedure should begin with the sample data normalization ranging from 0 to 1 employing Equation (21):

$$z = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (21)$$

where x_{\min} and x_{\max} are the minimal and maximal value of each import data series.

The sample data are classified into the training set and testing set. In this study, the last four data of CO₂ emissions and related driving factors are selected as the testing set. Therefore, the sample points of the training set range from 1991 to 2010, and the prediction points of the testing set begin in 2011 and end in 2014. The data of China's CO₂ emissions, GDP, energy consumption and population from 1991 to 2014 are indicated in Figure 4.

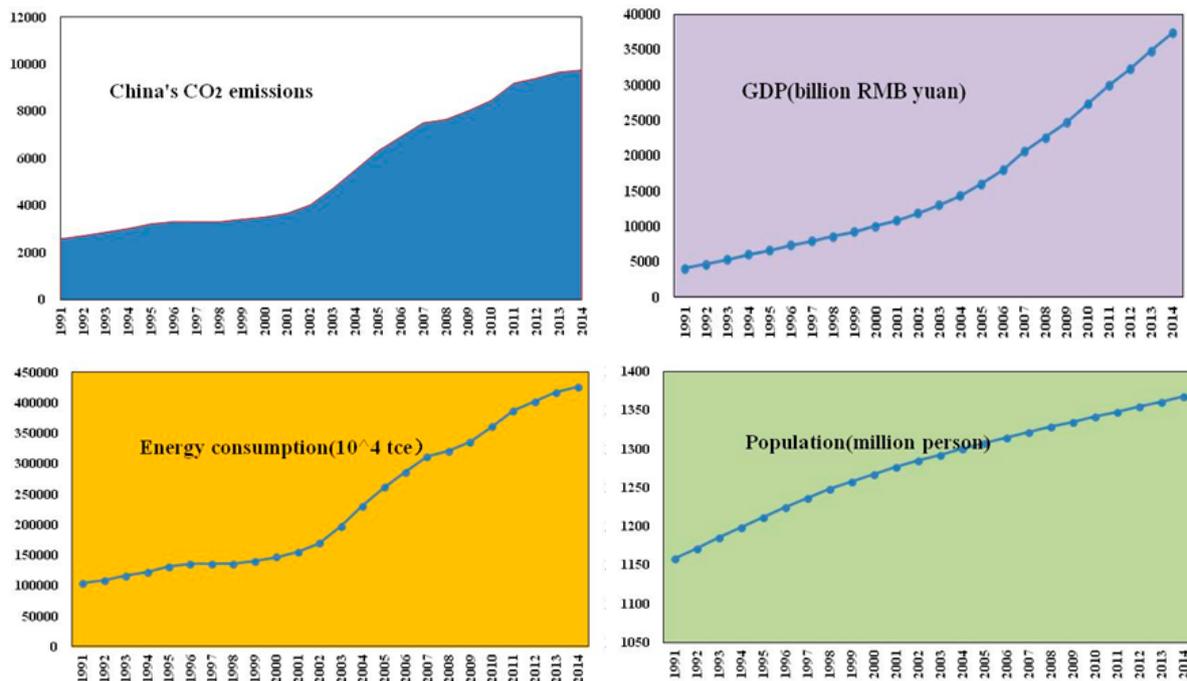


Figure 4. The sample data of China’s CO₂ emissions and selected driving factors.

3.2. Optimal Parameters Determination for LSSVM Method

In WOA-LSSVM method, the optimal values of “ σ^2 ” and “ c ” in LSSVM approach are determined by WOA method. Figure 5 displays the iterative root mean square error (RMSE) tendency of the WOA-LSSVM method finding for the optimized parameters. The iterative RMSE trend converges in the 6th generation after 100 iterations, and the optimal values of “ σ^2 ” and “ c ” are 2.0684 and 93.2203, respectively, which will be applied for CO₂ emissions prediction at the testing stage. Therefore, the CO₂ emissions during the testing period of 2011 to 2014 can be forecasted, and the prediction data are displayed in Table 2. As illustrated in Table 2, the gaps between actual CO₂ emissions values and forecasting CO₂ emissions values are small. Specifically, the gaps between actual and forecasting CO₂ emissions values at the year of 2011 and 2014 are only −3.55 million tons (Mt) and 7.71 Mt, respectively.

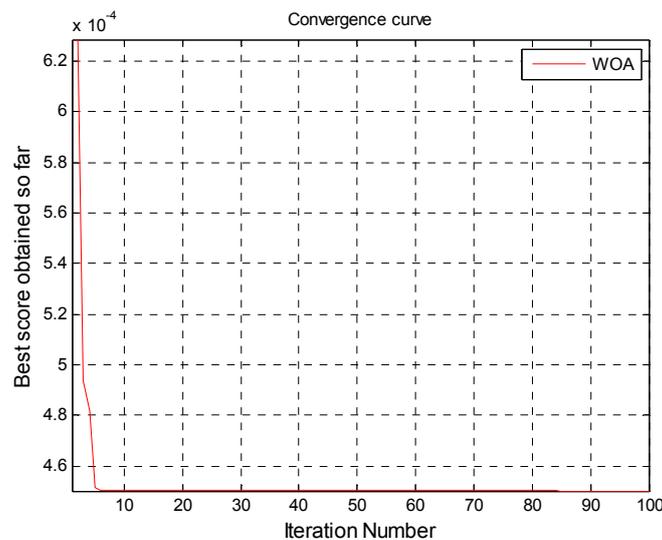


Figure 5. The iterative root mean square error (RMSE) trend of the WOA-LSSVM method finding for optimized parameters.

Table 2. Forecasting results and gaps by using WOA-LSSVM model.

| Year | Forecasting Value (Unit: Mt) | Actual Value (Unit: Mt) | The Gap * (Unit: Mt) |
|------|------------------------------|-------------------------|----------------------|
| 2011 | 9209.67 | 9206.12 | −3.55 |
| 2012 | 9378.71 | 9415.42 | 36.71 |
| 2013 | 9611.85 | 9674.22 | 62.37 |
| 2014 | 9753.36 | 9761.08 | 7.71 |

* The gap indicates the difference between the predicted value and actual value, the negative value of which demonstrates the actual data is less than the predicted one.

4. Forecasting Performance Evaluation

4.1. Selection of Comparison Models and Forecasting Performance Evaluation Index

For the purpose of evaluating the prediction accuracy of WOA-LSSVM method for China's CO₂ emissions forecasting, two things are essential to be determined, which are the choice of compared prediction methods, and the selection of the prediction performance evaluation indicator.

In order to comparatively analyze the prediction data of various methods, three compared prediction methods are chosen that are the LSSVM optimized by FOA (FOA-LSSVM), single LSSVM without parameters optimization, and OLS. For the FOA-LSSVM approach, the parameters " σ^2 " and " c " are optimized by FOA. The iteration begins with programming the original parameters of FOA: $maxgen = 100$, $sizepop = 20$, $(X_axis, Y_axis) \subset [-50, 50]$, $FR \subset [-10, 10]$. For the OLS model, through minimizing the residual sum of squares, the OLS model can calculate the coefficients of a linear model [33]. For these selected models, just like that of WOA-LSSVM model, the training sample ranges from 1991 to 2010, and the data between 2011 and 2014 are treated as the testing sample. Meanwhile, the driving factors including GDP, energy consumption and population are also regarded as the input variables of three compared prediction methods, and the output variable is CO₂ emissions.

The determined parameters " σ^2 " and " c " of the FOA-LSSVM model, the single LSSVM model and the WOA-LSSVM model are all illustrated in Table 3.

Table 3. The determined values of parameters for fruit fly optimization (FOA)-LSSVM, single LSSVM model and WOA-LSSVM model.

| Model | FOA-LSSVM | LSSVM | WOA-LSSVM |
|------------|-----------|-------|-----------|
| σ^2 | 0.2496 | 0.8 | 2.0684 |
| C | 5.6947 | 20 | 93.2203 |

To further analyze the prediction precision of each model, percentage error (PE), mean absolute percentage error (MAPE) and root mean square error (RMSE) are chosen to compare the prediction data of different models. They are computed by Equations (20)–(23), respectively.

$$PE = \frac{x(k) - \hat{x}(k)}{x(k)} \times 100\% \quad (22)$$

$$MAPE = \frac{1}{n} \sum_{k=1}^n \left| \frac{x(k) - \hat{x}(k)}{x(k)} \right| \times 100\% \quad (23)$$

where $x(k)$ represents the actual data at time k , and $\hat{x}(k)$ indicates the prediction data at time k .

4.2. Comparisons of Prediction Performance for Different Prediction Methods

The actual data and prediction data for CO₂ emissions in China by using WOA-LSSVM, FOA-LSSVM, LSSVM, and OLS methods during the period of 2011–2014 are illustrated in Figure 6. We can deduce that the WOA-LSSVM method performs best for CO₂ emissions in 2011 and 2014

among all of the compared forecasting models. In 2013, the forecasted result attained from the OLS method has the smallest gap from the actual value, and the WOA-LSSVM model ranks second. In 2012, the forecasting result of FOA-LSSVM model is closest to the actual value, followed by OLS method, WOA-LSSVM, and single LSSVM models.

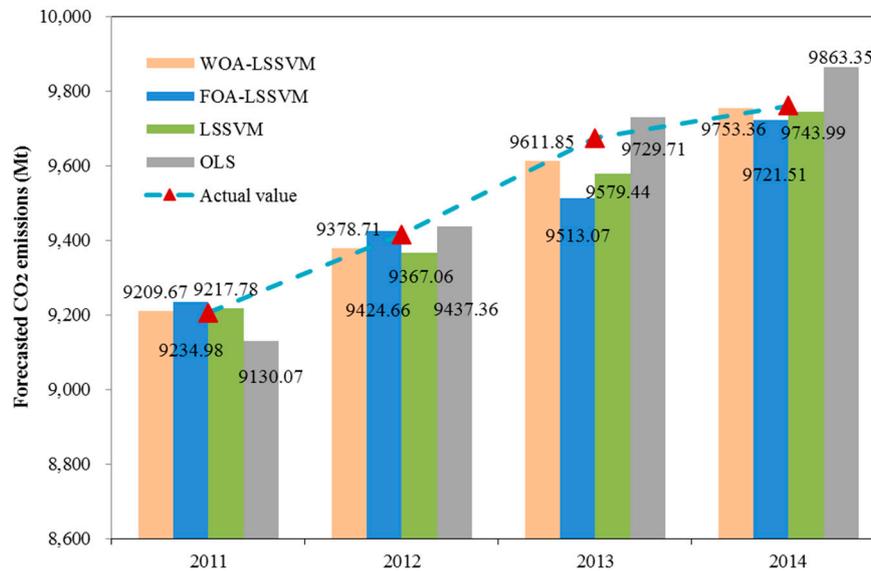


Figure 6. Forecasting results of China's CO₂ emissions between 2011 and 2014 by using different models.

To accurately compare the CO₂ emissions forecasting performances of different models, the evaluation indicators PE, MAPE, and RMSE are employed, and the results are demonstrated in Figure 7 and Table 4. Generally, the PE range in interval $[-3\%, 3\%]$ is deemed as standard performance of forecasting results [34]. As can be seen from Figure 7, all of the prediction points are in the range of $[-3\%, 3\%]$. Meanwhile, the forecasting PEs of WOA-LSSVM model all fall within the scope of $[-0.7\%, 0.7\%]$. The proposed WOA-LSSVM model show the best CO₂ emissions forecasting performance in 2011 and 2014 based on PE standard. Although the proposed WOA-LSSVM model does not reach the highest prediction accuracy in the remaining years based on PE standard, the gaps are small. Therefore, the WOA-LSSVM model has the superior CO₂ emissions prediction performance compared to OLS, single LSSVM, and FOA-LSSVM models.

The calculation results listed in Table 4 also verify the superiority of WOA-LSSVM model in accordance with MAPE and RMSE. The MAPE value of WOA-LSSVM model is only 0.29%, which is the smallest among all of the compared models (0.62% of FOA-LSSVM, 0.45% of single LSSVM, and 0.67% of OLS model). The RMSE value of WOA-LSSVM approach is significantly less than that calculated by FOA-LSSVM, LSSVM, and OLS methods, which are 36.43 Mt, 84.34 Mt, 54.20 Mt, and 70.36 Mt, respectively. In this case, the prediction performance of WOA-LSSVM model is better than FOA-LSSVM, single LSSVM, and OLS. It is also very interesting to discover that the prediction precision of single LSSVM method is better than the FOA-LSSVM, but that the WOA-LSSVM has the best prediction performance. This implies that the WOA has a greater capacity for optimizing LSSVM-based CO₂ emissions forecasting model than FOA. Therefore, it can be safely concluded that the WOA has great potential in optimizing parameters of LSSVM model for improving the CO₂ emissions forecasting accuracy.

Conclusively, the established WOA-LSSVM approach can dramatically lessen the deviations between the actual CO₂ emissions and prediction CO₂ emissions, which is significantly superior to the single LSSVM, FOA-LSSVM, and OLS models in China's CO₂ emissions forecasting. As a new kind of nature-inspired meta-heuristic optimization algorithm, the WOA has the prospect for wide application.



Figure 7. Percentage error (PE) values of selected prediction methods.

Table 4. Mean absolute percentage error (MAPE) and RMSE for chosen prediction methods.

| Model | WOA-LSSVM | FOA-LSSVM | LSSVM | OLS |
|-----------|-----------|-----------|-------|-------|
| MAPE (%) | 0.29 | 0.62 | 0.45 | 0.67 |
| RMSE (Mt) | 36.43 | 84.34 | 54.20 | 70.36 |

5. Conclusions

The accurate and reliable prediction on CO₂ emissions is of great significance for climate policy decision making and energy planning. In previous studies, it can be seen that the dramatic growth of CO₂ emissions in China is driven by many factors. However, the nonlinear relationships of CO₂ emissions with its driving forces are quite complicated, which makes the CO₂ emissions forecasting a difficult task. Therefore, improving the accuracy of CO₂ emissions forecasting is a field worth studying. In this study, the widely utilized LSSVM technique has been applied to build a new CO₂ emissions forecasting model, the parameters of which are optimized by a latest nature-inspired meta-heuristic optimization algorithm, WOA. Meanwhile, the proposed WOA-LSSVM model takes the main driving factors including GDP, energy consumption and population as the input variables, which takes the social economic effects on CO₂ emissions into consideration. To validate the potential of the presented WOA-LSSVM method in terms of CO₂ emissions forecasting, three other models (namely FOA-LSSVM, single LSSVM, and OLS) are selected to compare the forecasting performance. The comparative analysis implies that the proposed WOA-LSSVM method show the best performance in China's CO₂ emissions forecasting based on three evaluation indicators (namely PE, MAPE and RMSE). The proposed WOA-LSSVM model is significantly superior to other selected models in terms of CO₂ emissions forecasting, and the application of WOA to optimize LSSVM approach is valid and feasible. In future studies, the WOA, as a new nature-enlightened heuristic optimization algorithm, has the potential in optimizing other intelligent algorithms such as SVM and neural network for practical issues, such as electricity demand forecasting and renewable power output forecasting. In the meantime, we can also conduct research concentrating on scenario forecasting methods applied to CO₂ emissions forecasting to give full consideration to the impacts of national policies, such as carbon-reduction policies, economic and social development levels, and human behaviors on CO₂ emissions.

Acknowledgments: This paper is sponsored by the National Key R&D Program of China under Grant No. 2016YFB0900501, National Natural Science Foundation of China under Grant No. 71373076, the Fundamental Research Funds for the Central Universities under Grant No. 2017MS060 and 2017XS106.

Author Contributions: Huiru Zhao put forward the concept of this study. Haoran Zhao completed the paper. Sen Guo modified the final draft of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Suganthi, L.; Samuel, A.A. Energy models for demand forecasting—A review. *Renew. Sustain. Energy Rev.* **2012**, *16*, 1223–1240. [[CrossRef](#)]
2. Wang, W.; Kuang, Y.; Huang, N. Study on the decomposition of factors affecting energy-related carbon emissions in Guangdong province, China. *Energies* **2011**, *4*, 2249–2272. [[CrossRef](#)]
3. Akbostancı, E.; Tunç, G.I.; Aşık, S.T. CO₂ emissions of Turkish manufacturing industry: A decomposition analysis. *Appl. Energy* **2011**, *88*, 2273–2278. [[CrossRef](#)]
4. Auffhammer, M.; Carson, R.T. Forecasting the path of China's CO₂ emissions using province-level information. *J. Environ. Econ. Manag.* **2008**, *55*, 229–247. [[CrossRef](#)]
5. Safdarnejad, S.M.; Hedengren, J.D.; Baxter, L.L. Dynamic optimization of a hybrid system of energy-storing cryogenic carbon capture and a baseline power generation unit. *Appl. Energy* **2016**, *172*, 66–79. [[CrossRef](#)]
6. Gopan, A.; Kumfer, B.M.; Phillips, J.; Thimsen, D.; Smith, R.; Axelbaum, R.L. Process design and performance analysis of a staged, pressurized oxy-combustion (SPOC) power plant for carbon capture. *Appl. Energy* **2014**, *125*, 179–188. [[CrossRef](#)]
7. Safdarnejad, S.M.; Hedengren, J.D.; Baxter, L.L. Plant-level dynamic optimization of cryogenic carbon capture with conventional and renewable power sources. *Appl. Energy* **2015**, *149*, 354–366. [[CrossRef](#)]
8. Kang, C.A.; Brandt, A.R.; Durlinsky, L.J. A new carbon capture proxy model for optimizing the design and time-varying operation of a coal-natural gas power station. *Int. J. Greenh. Gas Control* **2016**, *48*, 234–252. [[CrossRef](#)]
9. Belaissaoui, B.; Cabot, G.; Cabot, M.S.; Willson, D.; Favre, E. CO₂ capture for gas turbines: An integrated energy-efficient process combining combustion in oxygen-enriched air, flue gas recirculation, and membrane separation. *Chem. Eng. Sci.* **2013**, *97*, 256–263. [[CrossRef](#)]
10. O'Neill, B.C.; Desai, M. Accuracy of past projections of US energy consumption. *Energy Policy* **2005**, *33*, 979–993. [[CrossRef](#)]
11. Auffhammer, M. The rationality of EIA forecasts under symmetric and asymmetric loss. *Resour. Energy Econ.* **2007**, *29*, 102–121. [[CrossRef](#)]
12. Bohringer, C.; Conrad, K.; Loschel, A. Carbon taxes and joint implementation: An applied general equilibrium analysis for Germany and India. *Environ. Resour. Econ.* **2003**, *24*, 49–76. [[CrossRef](#)]
13. Meng, M.; Niu, D. Modeling CO₂ emissions from fossil fuel combustion using the logistic equation. *Energy* **2011**, *36*, 3355–3359. [[CrossRef](#)]
14. Liang, Q.M.; Fan, Y.; Wei, Y.M. Multi-regional input–output model for regional energy requirements and CO₂ emissions in China. *Energy Policy* **2007**, *35*, 1685–1700. [[CrossRef](#)]
15. Chen, T.; Wang, Y. A fuzzy-neural approach for global CO₂ concentration forecasting. *Intell. Data Anal.* **2011**, *15*, 763–777.
16. Pao, H.T.; Tsai, C.H. Modeling and forecasting the CO₂ emissions, energy consumption, and economic growth in Brazil. *Energy* **2011**, *36*, 2450–2458. [[CrossRef](#)]
17. Lin, C.S.; Liou, F.M.; Huang, C.P. Grey forecasting model for CO₂ emissions: A Taiwan study. *Appl. Energy* **2011**, *88*, 3816–3820. [[CrossRef](#)]
18. He, J.K.; Deng, J.L.; Su, M.S. CO₂ emission from China's energy sector and strategy for its control. *Energy* **2010**, *35*, 4494–4498. [[CrossRef](#)]
19. Liao, R.J.; Zheng, H.B.; Grzybowski, S.; Yang, L.J. Particle swarm optimization-least squares support vector regression based forecasting model on dissolved gases in oil-filled power transformers. *Electr. Power Syst. Res.* **2011**, *81*, 2074–2080. [[CrossRef](#)]
20. Wu, Q. Hybrid model based on wavelet support vector machine and modified genetic algorithm penalizing Gaussian noises for power load forecasts. *Expert Syst. Appl.* **2011**, *38*, 379–385. [[CrossRef](#)]
21. Zhou, J.Y.; Shi, J.; Li, G. Fine tuning support vector machines for short-term wind speed forecasting. *Energy Convers. Manag.* **2011**, *52*, 1990–1998. [[CrossRef](#)]
22. Wang, S.; Yu, L.; Tang, L.; Wang, S.Y. A novel seasonal decomposition based least squares support vector regression ensemble learning approach for hydropower consumption forecasting in China. *Energy* **2011**, *36*, 6542–6554. [[CrossRef](#)]

23. Sulaimana, M.H.; Mustafab, M.W.; Shareefc, H.; Abd-Khalid, S.N. An application of artificial bee colony algorithm with least squares supports vector machine for real and reactive power tracing in deregulated power system. *Int. J. Electr. Power* **2012**, *37*, 67–77. [[CrossRef](#)]
24. Dos Santosa, G.S.; Justi Luvizottob, L.G.; Marianib, V.C.; Dos Santos, L.C. Least squares support vector machines with tuning based on chaotic differential evolution approach applied to the identification of a thermal process. *Expert Syst. Appl.* **2012**, *39*, 4805–4812. [[CrossRef](#)]
25. Li, H.; Guo, S.; Zhao, H.; Su, C.; Wang, B. Annual electric load forecasting by a least squares support vector machine with a fruit fly optimization algorithm. *Energies* **2012**, *5*, 4430–4445. [[CrossRef](#)]
26. Mirjalili, S.; Lewis, A. The Whale optimization algorithm. *Adv. Eng. Softw.* **2016**, *95*, 51–67. [[CrossRef](#)]
27. Huang, X.; Shi, L.; Suykens, J.A.K. Asymmetric least squares support vector machine classifiers. *Comput. Stat Data Anal.* **2014**, *70*, 395–405. [[CrossRef](#)]
28. Watkins, W.A.; Schevill, W.E. Aerial observation of feeding behavior in four baleen whales: *Eubalaena glacialis*, *Balaenoptera borealis*, *Megaptera novaeangliae*, and *Balaenoptera physalus*. *J. Mammal.* **1979**, *60*, 155–163. [[CrossRef](#)]
29. Goldbogen, J.A.; Friedlaender, A.S.; Calambokidis, J.; Mckenna, M.F.; Simon, M.; Nowacek, D.P. Integrative approaches to the study of baleen whale diving behavior, feeding performance, and foraging ecology. *BioScience* **2013**, *63*, 90–100.
30. Andresosso-O’Callaghan, B.; Yue, G. Sources of output change in China: 1987–1997 application of a structural decomposition analysis. *Appl. Econ.* **2002**, *34*, 2227–2237. [[CrossRef](#)]
31. Peters, G.; Webber, C.; Guan, D.; Hubacek, K. China’s growing CO₂ emissions a race between lifestyle changes and efficiency gains. *Environ. Sci. Technol.* **2007**, *41*, 5939–5944. [[CrossRef](#)]
32. Guan, D.; Peters, G.P.; Weber, C.L.; Hubacek, K. Journey to world top emitter: An analysis of the driving forces of China’s recent CO₂ emissions surge. *Geophys. Res. Lett.* **2009**, *36*, 1–5. [[CrossRef](#)]
33. Wang, S.; Yang, J. A probabilistic model for latent least squares regression. *Neurocomputing* **2015**, *149*, 1155–1161. [[CrossRef](#)]
34. Amiri, M.; Davande, H.; Sadeghian, A.; Chartier, S. Feedback associative memory based on a new hybrid model of generalized regression and self-feedback neural networks. *Neural Netw.* **2010**, *23*, 892–904. [[CrossRef](#)] [[PubMed](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).